

PlantSEED Workshop

In-depth Analysis of Genome Annotation

Viewing a User's Annotation

Genome TAIR10

Features

Annotations

Search annotations

1-10 of 653 results

< prev

next >

PlantSEED Role	Features	Exemplar Hits
Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	ATMG01360.1	ATMG01360.1
NADH dehydrogenase [ubiquinone] (EC 1.6.5.3) mitochondrial-encoded chain 2, ND2	ATMG00285.1	AT2G07689.1 ATCG00890.1 ATCG01010.1 ATCG01050.1 ATCG01250.1 ATMG00285.1 ATMG00513.1 ATMG00580.1
photosystem I iron-sulfur center subunit VII (PsaC)	ATCG01060.1	ATCG01060.1
Cytochrome b6-f complex subunit IV (PetD)	ATCG00730.1	ATCG00730.1
Cytochrome b6-f complex subunit, cytochrome b6	ATCG00720.1	AT2G07718.1 AT2G07727.1 ATCG00720.1 ATMG00220.1 ATMG00590.1
Photosystem II 10 kDa phosphoprotein (PsbH)	ATCG00710.1	ATCG00710.1
Photosystem II protein PsbN	ATCG00700.1	ATCG00700.1

Viewing a Feature

Genome TAIR10

Features

Annotations

Search annotations

1-10 of 653 results

< prev

next >

PlantSEED Role	Features	Exemplar Hits
Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	ATMG01360.1	ATMG01360.1
NADH dehydrogenase [ubiquinone] (EC 1.6.5.3) mitochondrial-encoded chain 2, ND2	ATMG00285.1	AT2G07689.1 ATCG00890.1 ATCG01010.1 ATCG01050.1 ATCG01250.1 ATMG00285.1 ATMG00513.1 ATMG00580.1
photosystem I iron-sulfur center subunit VII (PsaC)	ATCG01060.1	ATCG01060.1
Cytochrome b6-f complex subunit IV (PetD)	ATCG00730.1	ATCG00730.1
Cytochrome b6-f complex subunit, cytochrome b6	ATCG00720.1	AT2G07718.1 AT2G07727.1 ATCG00720.1 ATMG00220.1 ATMG00590.1
Photosystem II 10 kDa phosphoprotein (PsbH)	ATCG00710.1	ATCG00710.1
Photosystem II protein PsbN	ATCG00700.1	ATCG00700.1

Viewing a Feature

Genome Feature GRMZM2G176546

Function

Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3) [edit](#)

Subsystems

Fatty_acid_biosynthesis_in_plants_(plastidial)
Polyunsaturated_fatty_acids_biosynthesis_in_plants

Aliases

SEED	fig 381124.6.peg.36719
transcript	GRMZM2G176546_T01

Protein Sequence

```
MDDLAALRPTIFASVPRLYNRIYTAITNAVKESGGLKEKLFHTAYNAKRQAILKGKNPSVWDKLVFNKIKARLGGVRVRLMSSGASPLSADVMEFLRICFGGEVLEGYGMTETSCIIISAMDIGDRSIGHVGSPTASCEV  
KLVDPPEMNYTSDDQPYPRGEICVRGPIIFQGYKDEVQKKEVIDEDGWLHTGDIGLWLPGGRLKIIDRKNIFKLAQGEYIAPEKIENVYAKCKFIAQCFIYGDSFNSSLVGVAVEPEVMKAWAASEGIQCEDLRQL  
CADPRARAAVLADMDSVGKEAQLRGFEFAKAVRLVAEPFTVDNGLLTPTFKVQRQAKTYFAKEISDMYAEAREAEAPRSKL
```

Viewing a Feature

Plant Similarities

Prokaryotic Similarities

🔍 Search plant similarities

1-10 of 10 results

< prev next >

Hit ID	Genome	Function	Percent ID
GRMZM2G104847	Zmays-AGPv2	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
CHLREDRAFT_123147	Creinhardtii-v3.0	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
fgenes2_kg.6__2704__AT5G27600.1	Alyrata-v.1.0	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
GLYMA20G28200	Gmax-V1.0	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
BRADI1G48780	Bdistachyon-v1.0		79.71
GLYMA10G39540	Gmax-V1.0	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
GLYMA12G22220	Gmax-V1.0		79.71
fgenes2_kg.3__595__AT3G05970.1	Alyrata-v.1.0	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
GRMZM2G002614	Zmays-AGPv2	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71
Sb08g015332	Sbicolor-Sorbi1	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	79.71

Viewing a Feature

Plant Similarities

Prokaryotic Similarities

Search prokaryotic similarities

1-10 of 10 results

< prev next >

Hit ID	Genome	Function	Percent ID
fig 246197.24.peg.2487	Myxococcus xanthus DK 1622	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	29.52
fig 1150626.3.peg.253	Phaeospirillum molischianum	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	31.66
fig 762570.4.peg.2270	Rhodothermus marinus SG0.5JP17-172	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	32.07
fig 676032.3.peg.313	Francisella cf. novicida 3523	ATP-dependent DNA helicase RecG (EC 3.6.1.-)	32.36
fig 1141663.3.peg.3251	Providencia rettgeri Dmel1	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	29.68
fig 101510.15.peg.3763	Rhodococcus jostii RHA1	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	29.78
fig 1245027.3.peg.173	Proteus mirabilis C05028	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	30.73
fig 1114960.4.peg.3479	Rhodococcus pyridinivorans AK37	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	29.46
fig 991935.7.peg.3374	Vibrio cholerae HC-49A2	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	31.02
fig 991993.3.peg.3418	Vibrio cholerae CP1040(13)	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	31.02

Assessing Annotation

(True positives)

- One means by which one can evaluate annotation assigned using Kmers is by using BLAST
- The “Exemplar Hits” column contains the list of features which had a BLAST ‘hit’ with a curated Arabidopsis gene
- We can consider the annotation of a feature, via a K-mer, to be good, if there is a match in the “Exemplar Hits” column.

Assessing Annotation

(False negatives)

- A false negative occurs if a PlantSEED Functional Role was not assigned to any features, but there are “Exemplar Hits”.
- You may browse the “Exemplar Hits” and look at the Plant similarities for each feature.
- If you believe that the functional role can be assigned to any of these features, you can edit a feature’s function

Assessing Annotation (False negatives)

Genome Feature ATMG01170

Function

No function present

[edit](#)

Genome Feature ATMG01170

Function

4-hydroxybenzoate polyprenyltransferase (EC 2.5.1.39)

cancel [save](#)

Genome Feature ATMG01170

Function

4-hydroxybenzoate polyprenyltransferase (EC 2.5.1.39)

saving...

Assessing Annotation

(False positives)

- A false positive occurs if a PlantSEED Functional Role was assigned to any features, but there are zero “Exemplar Hits”.
- You may look at the Plant similarities for these features.
- If you believe that the functional role should not belong to the feature, you may edit the function so that it is an empty string.

Assessing Annotation

(Tasks)

- Look through every functional role in Central Carbon Metabolism
 - Glycolysis, TCA Cycle, PPP
- Does your genome have feature(s) annotated with each function?
- For missing roles, can you find and edit features?
- Tip: Use two tabs/pages
 - One for master list of PlantSEED annotation
 - One for annotation of User's genome